

**IN THE SPECIFICATION**

Please amend the specification as shown:

Please delete the paragraph on page 9, line 5 to page 10, line 5 and replace it with the following paragraph:

The vector can be any useful vector known to the ordinary artisan, including, but not limited to, a cloning vector, an insertion vector, or an expression vector. Examples of vectors include plasmids, phages, cosmids, phagemid, yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), human artificial chromosome (HAC), viral vector, such as adenoviral vector, retroviral vector, and other DNA sequences which are able to replicate or to be replicated *in vitro* or in a host cell, or to convey a desired DNA segment to a desired location within a host cell.

According to a preferred embodiment of the invention, the recombinant vector is a BAC pBeloBAC11 in which the genomic region of *Mycobacterium bovis-BCG* 1173P3 that spans the region corresponding to the locus 1,760,753 bp to 1,830,364 bp in the genome of *M. tuberculosis* H37Rv has been inserted into the HindIII restriction site; this recombinant vector is named X229. In this region, the inventors have demonstrated the deletion of a 2153 bp fragment, corresponding to SEQ ID N°4, in the vast majority of *M. tuberculosis* strains excepted strains of *M. tuberculosis* having the sequence CTG at codon 463 of gene *katG* and having no or very few IS6110 sequences inserted in their genome. That's the reason why the inventors named this deletion of 2153 bp TbD1 ("*M. tuberculosis* specific deletion 1"). TbD1 is flanked by the sequence GGC CTG GTC AAA CGC GGC TGG ATG CTG **(SEQ ID NO: 23)** and AGA TCC GTC TTT GAC ACG ATC GAC G **(SEQ ID NO: 24)**. External primers hybridizing with such sequences outside

TbD1 or the complementary sequences thereof can be used for the amplification of TbD1 to check for the presence or the absence of the deletion of the TbD1. The inventors design for example the following primers:

5'- CTA CCT CAT CTT CCG GTC CA-3' (SEQ ID N°17)

5'- CAT AGA TCC CGG ACA TGG TG-3' (SEQ ID N°18)

In order to get a specific 500 pb probe for hybridization experiments, a PCR amplification of a fragment comprised in TbD1 may be realized by using the plasmid X229 as a matrix. The amplification of a fragment of approximately 500 bp contained in TbD1 can be performed by using the following primers:

5'- CGT TCA ACC CCA AAC AGG TA-3' (SEQ ID N°13)

5'- AAT CGA ACT CGT GGA ACA CC-3' (SEQ ID N°14)

The amplification of a fragment of approximately 2,000 bp contained in TbD1 can be performed by using the following primers:

5'- ATT CAG CGT CTA TCG GTT GC-3' (SEQ ID N°15)

5'- AGC AGC TCG GGA TAT CGT AG-3' (SEQ ID N°16)

The PCR conditions are the following: denaturation 95°C 1 min, then 35 cycles of amplification [95°C during 30 seconds, 58°C during 1 min] , then elongation 72°C during 4 min.

Please delete the paragraph on page 20, lines 26-31 and replace it with the following paragraph:

**Figure 2** : Sequences **(SEQ ID NOS 116-117)** in the TbD1 region obtained from strains of various geographic regions.

\* refers to groups based on *katG*<sup>c463</sup>/*gyrA*<sup>c95</sup> sequence polymorphism defined by Sreevatsan and colleagues (Ref. 2). Numbers correspond to strain designation used in

Kremer et al. (1999, J. Clin Microbiol. 37: 2607-2618) (Ref. 8) and Supply et al (2001, J. Clin. Microbiol. 39: 3563-3571) (ref.9).

Please delete the paragraph on page 21, line 33 to page 22, line 4 and replace it with the following paragraph:

**Figure 6** : Sequence (**SEQ ID NO: 19**) of the specific insertion element in genome of *Mycobacterium canettii* strains. The beginning of this insertion element is at position 399 and the end of this insertion element is at position 2378. This insertion element contains the coding sequence of a putative transposase (sequence in bold characters, from position 517 to position 2307) that shows significant homology with a transposase of *Mycobacterium smegmatis*. This coding sequence is framed by two 20 bp inverted repeats (sequences underlined from position 399 to 418 and from position 2359 to 2378).

Please delete Table 1 on pages 36-39 and replace it with the following Table:

Table 1: RD, RvD and TbD1 regions and selected primers

Region	<u>Gene</u>	<u>Size</u>	<u>Internal Primerpair</u>	Flanking primers or
absent		<u>(kb)</u>		2 <sup>nd</sup> internal *
from BCG				primerpair

RD1	Rv3871- Rv3879c	9.5	RD1in-Rv3878F	RD1-flank.left
			GTC AGC CAA GTC	GAA ACA GTC CCC
			AGG CTA CC	AGC AGG T ( <u>SEQ ID</u>
			<u>(SEQ ID NO: 25)</u>	<u>NO: 26)</u>
			RD1in-Rv3878R	RD1-flank.right
			CAA CGT TGT GGT	TTC AAC GGG TTA
			TGT TGA GG	CTG CGA AT ( <u>SEQ</u>
			<u>(SEQ ID NO: 27)</u>	<u>ID NO: 28)</u>
RD2	Rv1978-Rv1988	10.8	RD2-Rv1979.int.F	RD2-flank.F
			TAT AGC TCT CGG	CTC GAC CGC GAC
			CAG GTT CC	GAT GTG C
			<u>(SEQ ID NO: 29)</u>	<u>(SEQ ID NO: 30)</u>
			RD2-Rv1979-int.R	RD2-flank.R
			ATC GGC ATC TAT	CCT CGT TGT CAC
			GTC GGT GT	CGC GTA TG ( <u>SEQ</u>
			<u>(SEQ ID NO: 31)</u>	<u>ID NO: 32)</u>
RD3*	Rv1573- Rv1586c	9.2	RD3-Rv1586.int.F	RD3-int-REP.F
			TTA TCT TGG CGT	CTG ACG TCG TTG
			TGA CGA TG	TCG AGG TA* ( <u>SEQ</u>
			<u>(SEQ ID NO: 33)</u>	<u>ID NO: 34)</u>

			RD3-Rv1586.int.R	RD3-int-REP.R
			CAT ATA AGG GTG	GTA CCC CCA GGC
			CCC GCT AC	GAT CTT*
			<u>(SEQ ID NO: 35)</u>	<u>(SEQ ID NO: 36)</u>
RD4	Rv1505c-	12.7	RD4-Rv1516.int.F	RD4-flank.F
	Rv1516c		CAA GGG GTA TGA	CTC GTC GAA GGC
			GGT TCA CG	CAC TAA AG <u>(SEQ</u>
			<u>(SEQ ID NO: 37)</u>	<u>ID NO: 38)</u>
			RD4-Rv1516.int.R	RD4-flank.R
			CGG TGA TTC GTG	AAG GCG AAC AGA
			ATT GAA CA	TTC AGC AT <u>(SEQ ID</u>
			<u>(SEQ ID NO: 39)</u>	<u>NO: 40)</u>
<b>Table 1</b>	<b>(continued)</b>			
RD5*	Rv2346c-	9.0	RD5A-Rv2348.int.F	RD5B-plcA.int.F
	Rv2353c		AAT CAC GCT GCT	CAA GTT GGG TCT
			GCT ACT CC	GGT CGA AT <u>(SEQ</u>
			<u>(SEQ ID NO: 41)</u>	<u>ID NO: 42)</u>
			RD5A-Rv2348.int.R	RD5B-plcA.int.R
			GTG CTT TTG CCT CTT	GCT ACC CAA GGT
			GGT C	CTC CTG GT <u>(SEQ</u>
			<u>(SEQ ID NO: 43)</u>	<u>ID NO: 44)</u>

RD6*	Rv3425-	4.9	RD6-IS1532F	ND
	Rv3428c		CAG CTG GTG AGT	
			TCA AAT GC	
			<b><u>(SEQ ID NO: 45)</u></b>	
			RD6-IS1532R	ND
			CTC CCG ACA CCT	
			GTT CGT	
			<b><u>(SEQ ID NO: 46)</u></b>	
RD7	Rv1964-Rv1977	12.7	RD7-Rv1976.int.F	RD7-flank.F
			TGG ATT GTC GAC	GGT AAT CGT GGC
			GGT ATG AA	CGA CAA G <b><u>(SEQ ID</u></b>
			<b><u>(SEQ ID NO: 47)</u></b>	<b><u>NO: 48)</u></b>
			RD7-Rv1976.int.R	RD7-flank.R
			GGT CGA TAA GGT	CAG CTC TTC CCC
			CAC GGA AC	TCT CGA C
			<b><u>(SEQ ID NO: 49)</u></b>	<b><u>(SEQ ID NO: 50)</u></b>
RD8	<i>ephA-lpqG</i>	5.9	RD8-ephA.F	RD8-flank.F
			GGT GTG ATT TGG	CAA TCA GGG CTG
			TGA GAC GAT G <b><u>(SEQ</u></b>	TGC TAA CC <b><u>(SEQ</u></b>
			<b><u>ID NO: 51)</u></b>	<b><u>ID NO: 52)</u></b>

			RD8-ephA.R	RD8-flank.R
			AGT TCC TCC TGA	CGA CAG TTG TGC
			CTA ATC CAG GC	GTA CTG GT <u>(SEQ</u>
			<u>(SEQ ID NO: 53)</u>	<u>ID NO: 54)</u>
RD9	<i>cobL</i> -Rv2075	2.0	RD9-intF	RD9-flankF
			CGA TGG TCA ACA	GTG TAG GTC AGC
			CCA CTA CG	CCC ATC C
			<u>(SEQ ID NO: 55)</u>	<u>(SEQ ID NO: 56)</u>
			RD9-intR	RD9-flankR
			CTG GAC CTC GAT	GCC CAA CAG CTC
			GAC CAC TC	GAC ATC
			<u>(SEQ ID NO: 57)</u>	<u>(SEQ ID NO: 58)</u>
RD10	Rv0221-Rv0223	1.9	RD10-intF	RD10-flankF
			GTA ACC GCT TCA	CTG CAA CCA TCC
			CCG GAA T	GGT ACA C
			<u>(SEQ ID NO: 59)</u>	<u>(SEQ ID NO: 60)</u>
			RD10-intR	RD10-flankR
			GTC AAC TCC ACG	GTC ATG AAC GCC
			GAA AGA CC	GGA CAG
			<u>(SEQ ID NO: 61)</u>	<u>(SEQ ID NO: 62)</u>

RD11	Rv2645-	11.0	RD11-Rv2646F	RD11-fla-F
	Rv2695c		CGG CAG CTA GAC	TCA CAT AGG GGC
			GAC CTC	TGC GAT AG <u>(SEQ</u>
			<u>(SEQ ID NO: 63)</u>	<u>ID NO: 64)</u>
			RD11-Rv2646R	RD11-fla-R
			AAC GTG CTG CGA	AGA GGA ACC TTT
			TAG GTT TT	CGG TGG TT <u>(SEQ</u>
			<u>(SEQ ID NO: 65)</u>	<u>ID NO: 66)</u>
RD12	sseC-Rv3121	2.8	RD12-Rv3120.int.F	RD12-flank.F
			GAA ATA CGA GTG	GCC ATC AAC GTC
			CGC TGA CC	AAG AAC CT <u>(SEQ</u>
			<u>(SEQ ID NO: 67)</u>	<u>ID NO: 68)</u>
			RD12-Rv3120.int.R	RD12-flank.R
			CTC TGA ACC ATC	CGG CCA GGT AAC
			GGT GTC G	AAG GAG T <u>(SEQ ID</u>
			<u>(SEQ ID NO: 69)</u>	<u>NO: 70)</u>
RD13	Rv1255c-	3.0	RD13intF	RD13-flank.F
	Rv1257c		GGA TGT CAC TCG	CGA TGG TGT TTC
			GAA CGG CA	TTG GTG AG <u>(SEQ</u>
			<u>(SEQ ID NO: 71)</u>	<u>ID NO: 72)</u>



			RD13intR	RD13-flank.R
			CAC CGG GCT GAT	GGA TCG GCT CAG
			CGA GCG A	TGA ATA CC <u>(SEQ ID</u>
			<u>(SEQ ID NO: 73)</u>	<u>NO: 74)</u>
RD14	Rv1765c-	9.0	RD14-Rv1769.int.F	RD14-flankF
	Rv1773c		GTG GAG CAC CTT	TTG ATT CGC CAA
			GAC CTG AT	CAA CTG AA <u>(SEQ</u>
			<u>(SEQ ID NO: 75)</u>	<u>ID NO: 76)</u>
			RD14-Rv1769.int.R	RD14-flankR
			CGT CGA ATA CGA	GGG CTG GTT AGT
			GTC GAA CA	GTC GAT TC <u>(SEQ</u>
			<u>(SEQ ID NO: 77)</u>	<u>ID NO: 78)</u>

**Table 1 (continued)**

**Region missing from *M. tuberculosis* H37Rv**

RvD1*	5.0	RvD1-int1F	RvD1-int2.F
		AGC GCG TCG AAC	GAG CCA CTC CGA
		ACC GGC	TGT TGA CT <u>(SEQ ID</u>
		<u>(SEQ ID NO: 79)</u>	<u>NO: 80)</u>
		RvD1-int1R	RvD1-int2.R
		CCT GAA TCC GCG	CAC GCG AAC CCT
		CAA TTC CAT	ACC TAC AT <u>(SEQ ID</u>
		<u>(SEQ ID NO: 81)</u>	<u>NO: 82)</u>

RvD2*	<i>plcD</i>	5.1	RvD2-int1F	RvD2-int2F
			GTT CTC CTG TCG	GGA CGG TGA CGG
			AAC CTC CA	TAT TTG TC <u>(SEQ ID</u>
			<u>(SEQ ID NO: 83)</u>	<u>NO: 84)</u>
			RvD2-int1R	RvD2-int2R
			ACT TCA CCG GTT	TCG CCA ACT TCT
			TCA TCT CG	ATG GAC CT <u>(SEQ</u>
			<u>(SEQ ID NO: 85)</u>	<u>ID NO: 86)</u>
RvD3		1.0	RvD3-intF	RvD3-flank.F
			ATC GAT CAG GTC	AAA CCA TGC AGC
			GTC AAT GC	GTC TGC CA <u>(SEQ</u>
			<u>(SEQ ID NO: 87)</u>	<u>ID NO: 88)</u>
			RvD3-intR	RvD3-flankR
			ACG CCA CCA TCA	GCG TTT CTG CGT
			AGA TCC	CTG GTT GA <u>(SEQ</u>
			<u>(SEQ ID NO: 89)</u>	<u>ID NO: 90)</u>
RvD4*	PPE gene	0.8	RvD4-intF-PPE	ND
			GGT TGC CAA CGT	
			TAC CGA TGC	
			<u>(SEQ ID NO: 91)</u>	

			RvD4-intR-PPE	ND
			CCG GTG GTG GTG	
			GCG GCT	
			<u>(SEQ ID NO: 92)</u>	
RvD5	<i>moa</i>	4.0	RvD5intF	RvD5-flankF
			GGG TTC ACG TTC	CCC ATC GTG GTC
			ATT ACT GTT C	GTT CAC C
			<u>(SEQ ID NO: 93)</u>	<u>(SEQ ID NO: 94)</u>
			RvD5intR	RvD5-flankR
			CCT GCG CTT ATC	GTA CCC GCA CCA
			TCT AGC GG	CCT GCT G
			<u>(SEQ ID NO: 95)</u>	<u>(SEQ ID NO: 96)</u>
TbD1	<i>mmpL6</i>	2.1	TbD1intS.F	TbD1fla1-F
			CGT TCA ACC CCA	CTA CCT CAT CTT
			AAC AGG TA	CCG GTC CA <u>(SEQ</u>
			<u>(SEQ ID NO: 97)</u>	<u>ID NO: 98)</u>
			TbD1intS.R	TbD1fla1-R
			AAT CGA ACT CGT	CAT AGA TCC CGG
			GGA ACA CC	ACA TGG TG <u>(SEQ</u>
			<u>(SEQ ID NO: 99)</u>	<u>ID NO: 100)</u>

*katG*, *gyrA*, *oxyR*', *pncA* and *mmpL6* PCR and sequencing primers

*katG*<sup>463</sup>

<i>katG</i> -2154,225-PCR-F	<i>katG</i> -2154,872-SEQ-
CTA CCA GCA CCG	R
TCA TCT CA	ACA AGC TGA TCC
<u>(SEQ ID NO: 101)</u>	ACC GAG AC <u>(SEQ</u>
	<u>ID NO: 102)</u>

*katG*-2155,157-PCR-R

AGG TCG TAT GGA

CGAACA CC

(SEQ ID NO: 103)

*gyrA*<sup>95</sup>

<i>gyrA</i> -7,127-PCR-F	<i>gyrA</i> -7,461F
GTT CGT GTG TTG	CGG GTG CTC TAT
CGT CAA GT	GCA ATG TT <u>(SEQ ID</u>
<u>(SEQ ID NO: 104)</u>	<u>NO: 105)</u>

*gyrA*- 8,312-PCR-R

CAG CTG GGT GTG

CTT GTA AA

(SEQ ID NO: 106)

*oxyR*<sup>285</sup>

<i>oxyR</i> 2725,559F	<i>oxyR</i> -2726,024-SEQ-
TAT GCG ATC AGG	R
CGT ACT TG	CAA AGC AGT GGT
<u>(SEQ ID NO: 107)</u>	TCA GCA GT <u>(SEQ</u>
	<u>ID NO: 108)</u>

*oxyR*-2726,024-PCR-R

CAA AGC AGT GGT

TCA GCA GT

**(SEQ ID NO: 109)**

**Table 1 (continued)**

*pncA*<sup>57</sup>

*pncA*-2288,678-PCR-F

*pncA*- 2289,319-SEQ-

ATC AGG AGC TGC

R

AAA CCA AC

GGC GTC ATG GAC

**(SEQ ID NO: 110)**

CCT ATA TC **(SEQ ID**

**NO: 111)**

*pncA*- 2289,319-PCR-R

GGC GTC ATG GAC

CCT ATA TC

**(SEQ ID NO: 112)**

*mmpL6*<sup>551</sup>

*mmpL*-seq5F

*mmpL*-seq5F

GTA TCA GAG GGA

GTA TCA GAG GGA

CCG AGC AG

CCG AGC AG **(SEQ**

**(SEQ ID NO: 113)**

**ID NO: 114)**

TBD1fla1-R

CAT AGA TCC CGG

ACA TGG TG

**(SEQ ID NO: 115)**

The RD nomenclature used in this table is based on that used by Brosch *et al.* (2000), (Ref. 25) and differs from that proposed by Behr and coworkers (1999), (Ref. 6). Primer sequences are shown in 5' → 3' direction.

\* Regions where a second pair of internal primers was used rather than flanking primers, due to flanking repetitive regions, and/or mobile genetic elements.